DOCUMENT RESUME

ED 324 952 FL 018 935

AUTHOR Van Elsen, Edwig; And Others
TITLE COMETT-CALLIOPE: The Implementation of Call Materials for Business and Industrial Purposes.
PUB DATE Mar 90
NOTE 23p.; Paper presented at the Annual Meeting of the Computer Assisted Language Learning and Instruction Consortium (Baltimore, MD, March 19-23, 1990).
PUB TYPE Speeches/Conference Papers (150) -- Reports - Descriptive (141)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Business Communication; *Computer Assisted Instruction; Computer Software Development; *Courseware; Data Collection; Foreign Countries; Instructional Materials; *Languages for Special Purposes; Learning Strategies; Program Descriptions; *Second Language Instruction; Second Language Learning; Teaching Methods; *Vocabulary
IDENTIFIERS *COMETT CALLIOPE (Project); European Community

ABSTRACT

The development of a Computer Assisted Language Learning for Information Organization and Production in Europe (CALLIOPE) program is discussed. CALLIOPE is a program launched by the European Community that is intended to provide computer-based foreign language instruction for the business and industrial environment. Program goals are two-fold: (1) to investigate specific terminology employed by participating companies; and (2) to include that terminology in a computer-based instructional program that affords efficient teaching of specialized second-language vocabularies. The development of a program providing relevant and made-to-measure courseware materials for one company is described in detail. Crucial to the program's success was a careful investigation of existing language needs in the company, the collection of company authentic data materials, intelligent analysis and coding of the materials for database use, and the formation of instructional strategies and user interfaces to increase the versatility and multifunctionality of the final program. (JL)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED324952

COMETT-CALLIOPE : The Implementation of Call Materials for Business and Industrial Purposes.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY
E. Van Elsen
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Edwig Van Elsen[1], Jozef Colpaert and Wilfried Decoo

0. Comett-Calliope in a few words

The *Comett programme*, launched by the European Community, aims to establish a close collaboration network between various European companies and universities with respect to training and education projects in the domain of technology.

In February 1987 *Didascalia*, the research team of prof. dr. W. Decoo at Antwerp University (U.I.A.) and K. Vervaeke, personnel manager of the *N.V. Bekaert* in Belgium, initiated Calliope as a Comett project. In June 1987 they submitted, together with another 10 companies and 6 universities from all over Europe, the project description to the E.C. The proposal was accepted in December 1987.

Calliope[2] aims at the development of foreign language courseware for the company and business environment. One important subpart of the project is the development of courseware with LSP[3] vocabulary contents. Calliope's goals are twofold: an investigation of participating companies' specific terminologies and the implementation of those contents in a CALL environment which allows efficient teaching and training of specialized terminology.

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
X This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

FL018935

We will elaborate on the various stages that constitute the development process of these programs. Findings will be illustrated with actual data from *'Becall'*, one of the Calliope programs, developed for the *N.V. Bekaert*.

1. Calliope Frameworks

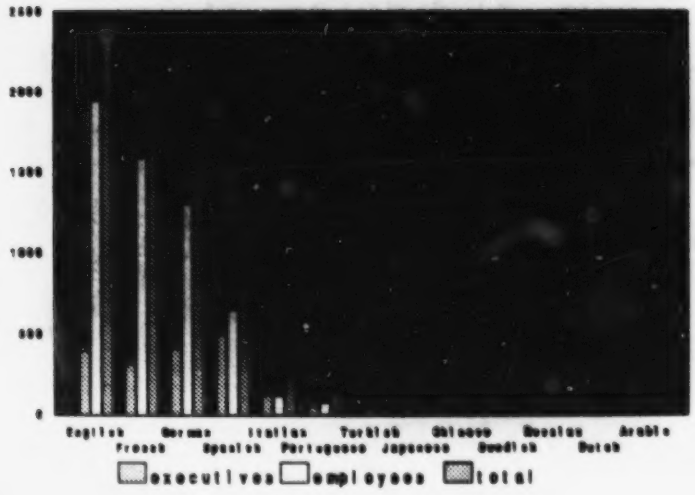
During the three years of EC-Comett funding, Calliope created a general consultancy plan which outlines in a detailed way the overall prerequisites, the necessary production steps, the time schedule and the necessary budget for LSP CALL courseware development. This consultancy plan (*'Calliope Frameworks'*) was based on the practical experience gained from two pilot projects: *'Becall'* (steelwire terminology) and *'Bellingua'* (telecom terminology), developed in collaboration with *Bekaert* en *Alcatel-Bell* respectively. Calliope Frameworks includes 4 main development stages: (1) a number of preparatory stages (2) the definition of the content base layout (3) a courseware engineering phase and (4) the implementation of the developed teaching materials in the existing training infrastructure of the companies involved. Since the two pilot projects have only recently entered the final stage, we will here comment on the first three stages only.

2. Preparatory stages

2.1. Analysis of the existing language needs

A first condition for efficient courseware development is a precise answer to the question 'what are we going to teach to whom?'. Are we going to focus exclusively on English and German as a foreign language or will we include for instance Spanish and Italian too? Would it make sense to include juridical terminology or should we stick to the words frequently used in the production department only? Will such a program merely be used by the executives of the company or will the entire staff be interested? Intuitively we thought that these and similar questions would be easy for the company management to answer. In practice however, the answers to our preliminary questions turned out to be most divergent and scarcely ever based on objective data.

It was therefore decided to include a language needs survey in Calliope Frameworks: a simple but straightforward questionnaire to be filled in by all staff members, combined with an automated analysis tool, allowed us to accurately and rapidly chart the language needs landscape of participating companies. The survey for Bekaert for instance revealed that there was a highly predominant need for English, French, German and Spanish as a foreign language, that there was a small but still significant need for Italian and Portuguese, but that no other languages seemed relevant in the Bekaert situation.



Another major relevant conclusion from the language survey organized in Bekaert, was that the relative importance for the 4 major languages differed for executives and employees: Spanish and German were considered more relevant by the executives than by the employees, who in contrast thought training in English and French to be more valuable for their working efficiency.

2.2. Definition of target public and end user

Given the information from the automated survey, the next step was to define the eventual target public for the courseware materials under development. In the Bekaert context, it was clear that both the likely end users and their needs were extremely heterogeneous. There was a majority of employees interested in a wide range of technical fields (production, R&D, finance, ...). As for the languages themselves, it was clear that prior attention had to be given to English, French, German and Spanish. Though there was some potential for Italian and Portuguese

on a long-term basis, an investment in CALL software for the other languages mentioned in the survey proved clearly out of order.

Calliope could in the Bekaert context not do with a fixed end user / fixed content program, specifically designed for, for instance, the unskilled workers in the production departments of Spanish subsidiaries. There was on the contrary a strong need for a highly flexible program, where extensive content materials (6 languages and as many lexical domains as there are departments in the company) were made easily accessible and selectable for the end user. In section 4.2.1 we will outline in which way this flexibility criterion was met.

2.3. Collecting authentic data

The majority of our informants stressed the importance for the language in the final program to match as closely as possible the language actually being spoken and written in the company. The frustration of most past language courses was exactly situated in their non being adapted to the language actually used in the company involved. This demand for authentic contents gave rise to a rigorously frequency based approach in Calliope.

A corpus was generated consisting of up-to-date texts, written, read and/or used at the departments where the survey had indicated a clear interest and need in foreign language training. This corpus included internal reports and articles, user manuals, technical documentation, journal articles, business correspondence, etc. In some cases the training department thought it useful to combine these data with

standardized content material, i.e. glossaries with standardized terminology defined by official bodies such as National Councils, the EC, etc.

2.4. Processing the data materials

The raw materials collected under 2.3. were next to be scanned into machine readable format. Several corpus analysis tools, especially developed by the Calliope project, made an intelligent data analysis possible: words were transformed into headword presentation (infinitive, singular, masculine forms), texts were analyzed according to frequency values, frequent word combinations were automatically retrieved so that compounds and stereotyped expressions could be detected, etc. The final output of these processes were a number of frequency lists (based on several text corpora) which we had to combine in an automatic, but semicontrolled way. Frequency lists can differ in quality and importance for many reasons: the quantity of the corpora they are based on, the degree of idiosyncrasy of the texts they are based on, the type of texts they are based on, etc. The project therefore developed a program which allows the researcher to combine various frequency lists in a calibrated way, ie. s/he can indicate the relative importance of the lists included. The program will then combine and define the unified list according to the values given.

3. Defining the content base layout

3.1. Didactic formatting of the contents

The preparatory stages described in section 2 transformed a number of raw text corpora in one final list, which includes the most frequently used phrases, words and idioms of a participating company. This list will be the foundation for the database of the final courseware product. Every single item should now be defined and encoded so that it afterwards becomes readily available and accessible to the end user.

The rationale behind the Calliope approach is that every item is considered as a '*didacteme*': a learnable unit defined by a unique set of distinctive features. The properties which make an item like 'corrosive' an item that is worthwhile learning, are characteristics like it belonging to a particular semantic or technical field, it belonging to a specific word class, its translation, its frequency level and various linguistic features (gender, uncountable/countable, etc.). Information like this has somehow to be present in the database, as to guarantee that every entry in the content base gets maximal 'learnability'.

In short, every entry in the database was stored as a fixed length record structure with 7 fields of variable length: *(1) headword, (2) article, (3) context sentence, (4) frequency value, (5) semantic field value, (6) word class value and (7) linguistic characteristics.*

The latter four data were quantified and encoded as a four bytes data structure.

Considering a byte as a boolean value, we can store eight true/false indications (8 bits) in one byte. The alternative is to maximally exploit the byte as a numerical value between 0 and 255. In this way, we can store no less than 256 feature values (or distinctive features of a didacteme) in one byte. [4]

3.2. Development of learning and teaching strategies

A crucial factor for the final content base layout is the type and number of strategies you are going to include in the courseware. Doing multiple choice exercises requires a different type of content base structure from the type of database used for doing crosswords with the computer. Calliope consulted the training officers in the participating companies and asked them what in their experience were crucial criteria for efficient strategy design.

A first criterion they mentioned was that the program would need *a variety of exercises* with a growing degree of difficulty. This criterion was largely inspired by the fact that the program would serve a diversity of end users in the company, each working at his or her own level and pace with the vocabulary selected.

Secondly, it was argued that for adult education in a company setting there was a clear preference for the more *traditional types of exercises*. Though quite popular in CALL contexts, language games for instance were considered less appropriate for this specific learning situation.

Finally, it was stressed that the course should follow *a modular approach*, i.e. training officers indicated the need for structured and systematic progress in the

language training programs offered to the staff: one course should strictly build on and enlarge the knowledge acquired in another language course. Even course-internally, one of the biggest problems with language courses is that they often present their contents in a parallel way, where there is no systematic continuation and expansion of previously taught learning contents.

Calliope implemented these criteria into the courseware as follows: (a) end-users can train their selected vocabulary in 10 different strategies, ranging from a filling the gap exercise with hint via a written cloze exercise to the more traditional translation exercise (b) the core structure of all exercise types is the so-called *STS (Semantically Transparent Sentence) context*: every item is presented to the user in a context matching as closely as possible the natural contexts in which the item occurs under normal company situations. This context sentence is semantically transparent, i.e. it contains clear semantic indications which point to the item taught. Consequently, this item will be positioned as close to the end of the context as possible (for examples, see section 3.3). A crucial aspect is that these STS contexts only make use of terminology that from the frequency point of view is of a lower value than the item taught. In other words, the trainee is familiar with all semantic indicators in the sentence. So s/he will have little difficulty in understanding this context sentence and at the same time s/he will constantly and systematically rehearse previously acquired terminology. This methodology guarantees the modular approach from the program internal point of view: the contents are systematically organized.

3.3. Development of the content base

The next step is to write out the STS contexts and code clusters for all content base entries, according to the specifications outlined under 3.1. and 3.2. In the Bekaert program *'Becall'* , this was done for all 6 languages, every language constituting a different data file. The 6 language files were then linked in a parallel way, so that every language became available as either source or target language.

The following lines are some simplified extracts from the English BECALL content base. In the actual content base lines there is additional space for synonyms, semantic comments, gender indications, etc:

headword	article	STS context	Code cluster
creep	the	Stress at high temperatures will cause slow and continuous plastic deformation or *creep*.	17/2/140/000
pitting	-	After pickling, the acid may continue to be active and thus cause *pitting*.	18/2/152/100
hot dip	the	Steel wire is covered with molten zinc during the *hot dip*.	18/4/160/000

The STS contexts nicely illustrate the principle of semantic indication. Stress, high temperature, plastic deformation, pickling, acid, molten zinc, etc, are all semantic indicators, eliciting at the end of the phrase the term involved. The code clusters indicate between the slashes the frequency value of the headword, its word class, technical field and a variety of linguistic information respectively.

As for all STS contexts, it must be noted that their accuracy and their being

adapted to the actual company situation were continually tested and evaluated by a group of more than 50 test persons during trial runs and debug sessions in the company.

4. Courseware engineering

4.1. Data-conversion

In order to guarantee maximal compatibility with other systems and applications, the entire content base of 1500 entries and STS contexts in 6 languages was stored in 6 standard ASCII files of 230 Kb each. Given the fact that such a content base is far too extensive to allow rapid and efficient data input and output, these ASCII text data files were converted into compacted files of records.

4.2. User interface

4.2.1. External versatility

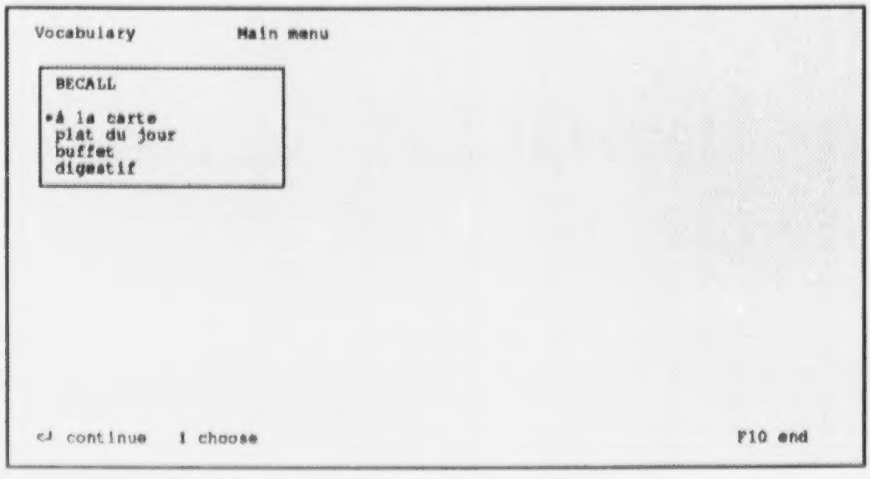
Multifunctionality was a decisive factor for our business partners to support a considerable and long-term investment in CALL materials. The programs developed should answer the needs of as many potential end users as possible: trainers, trainees, individual language students, trainer assisted language groups, executives, administrators, workers, etc.

External versatility is the key concept for all courseware materials developed by the Didascalia group[5]. An externally versatile program should be adaptable to the externally determined needs of its users. In a school situation the concept refers to the adaptability of courseware materials to a textbook, a pre-defined curriculum, a teacher or a student. In a company environment, externally versatile programs are adaptable to job profiles, time management systems and to the characteristics of existing training infrastructure. Program external factors in each case determine the exact nature of the end user's needs. Two main interface factors underlying the basic construction of the Calliope software realize this externally versatile design: a *selection system on extensive content bases* and *the multiple environment model.*

The idea of *a selection system* has implicitly been discussed in the previous sections: the idea is simple and straightforward: if you combine an extensive content base with a detailed and easy to handle selection system, there will be a considerable increase in the number of potential end users; all users will experience that the vocabulary which is relevant to their working situation is included in the program and that moreover they can precisely select their subvocabulary from a larger terminology bank so that they do not have to go through a lot of irrelevant language material. In the next section (4.2.2.) we will give a more detailed overview of the selection parameters present in this system.

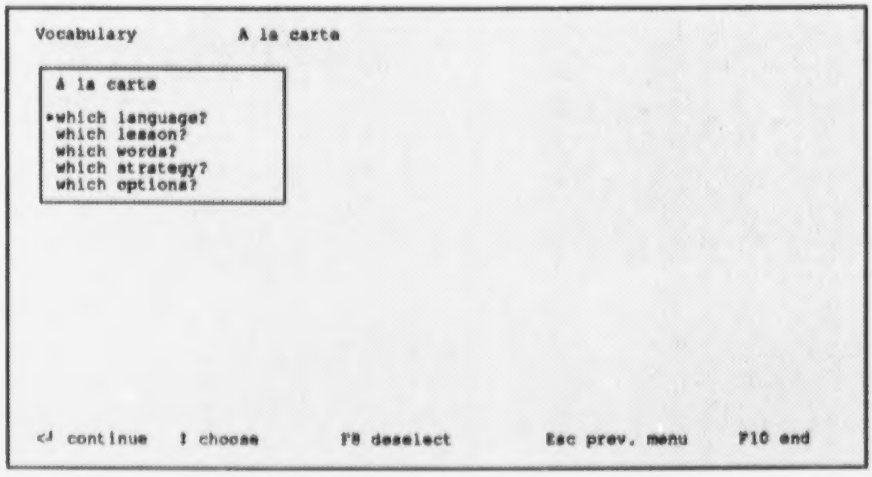
The *multiple environment* model makes the entire program fit for a multi-purpose use. The program includes four comparable but separate environments: one for more advanced individual learners, one for a trainer with a group of

trainees, one for error-analysis activities and one for reference activities. These separate environments are indicated by a set of culinary terms which are presented to the user in the *'main menu'* of the Calliope programs:



4.2.2. A la carte

'A la carte' is a typical environment for the more advanced user working individually with the program. This environment opens all the menus for the user who himself can select contents and strategies at will. S/He can do so by going through a series of convenient menus which appear in overlapping frames on the computer screen.

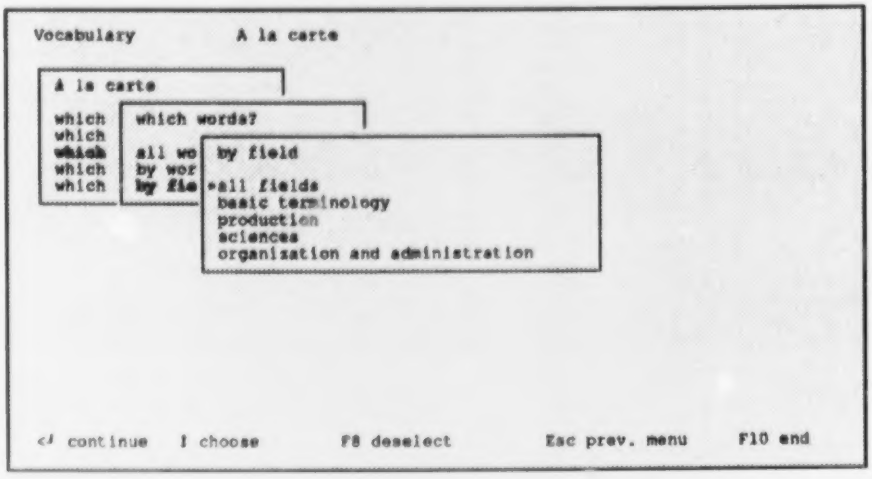


There are four actions to be taken in the 'à la carte' menu: (a) a user should select his target language in the *'which language?'* entry (his native language or the program's interface language has already been defined while booting the program) (b) in the *'which lesson?'* and *'which words?'* menus s/he should select the content material s/he is going to work with (c) a user should select particular strategies to actually train and practise the selected terminology under the *'which strategy'* menu (d) the user may, if s/he likes to, change some standard program settings (time limit, automatic typing error correction, sound, etc) in the *'which options?'* menu.

As for the *selection of contents* the user has at his or her disposal *3 parameters* for selecting items from the overall database. Items can be selected on the basis of their *frequency level*, the *word class* they belong to and the *semantic-technical field* they figure in.

As for the *frequency level*, the entire corpus has been subdivided in 6 levels of 250 items each, every level including again 5 units of 50 items. As such, trainees

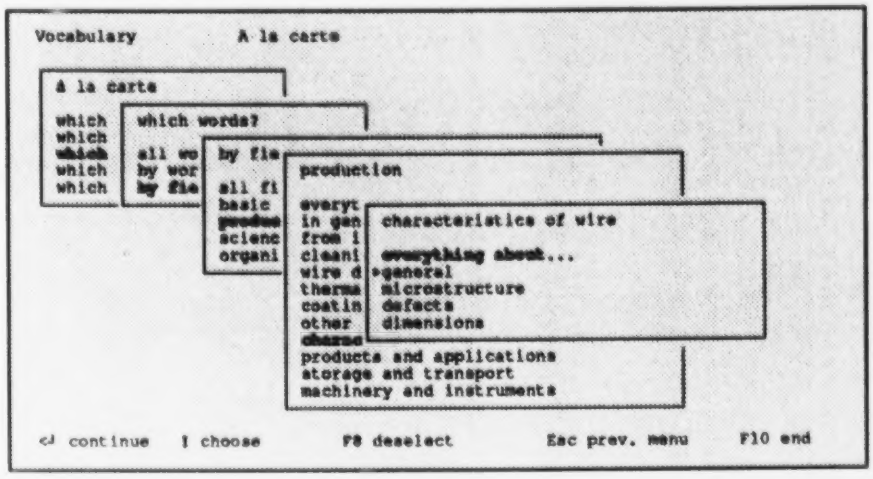
can go through the program in a strictly chronological way, lesson by lesson systematically extending their vocabulary knowledge, because the structure of the STS contexts guarantees they will only be confronted with words and expressions already acquired in previous lessons and levels.

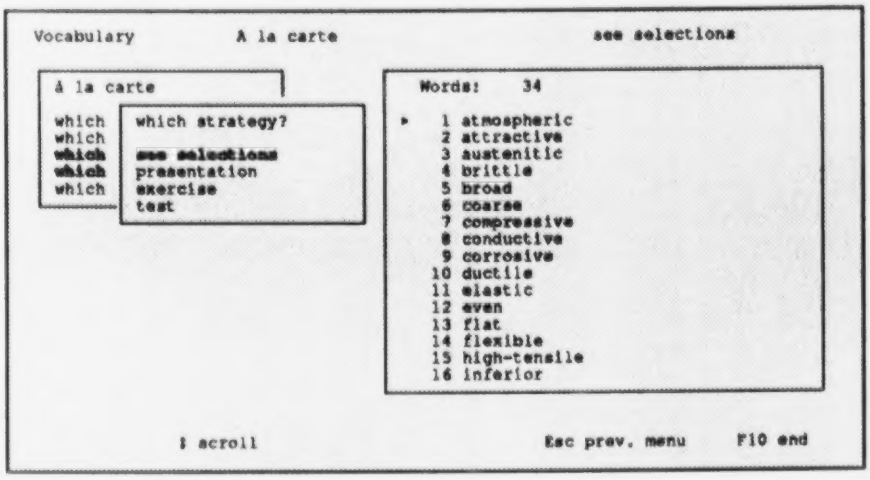


Secondly one can select terminology out of the 38 *semantic-technical fields* that are present in a tree-structured menu-system under the '*by field*' heading. Four main areas of Bekaert terminology are grouped under the labels 'basic terminology', 'production', 'sciences', and 'organization and administration'. All of these four entries open in their turn further submenus which guide the end user precisely to a specific subdomain of vocabulary that is relevant to his or her needs.

Thirdly, one can narrow down one's selection by defining a particular *word class* (nouns, adjectives, verbs or expressions). This selection parameter is most efficient when it is used in combination with the other selection criteria. It is indeed possible to combine at will all content selection parameters. In this way we can for instance select all adjectives which express characteristics of wire; a

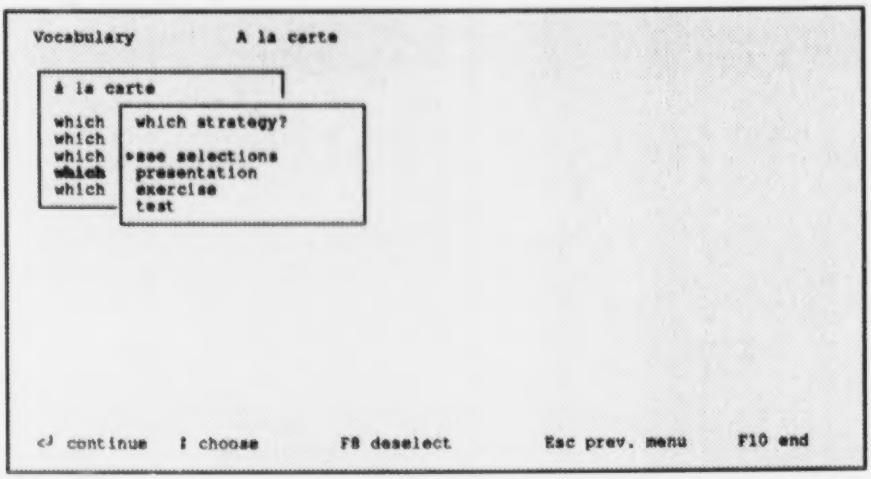
selection which yields a set of 34 items from the database.



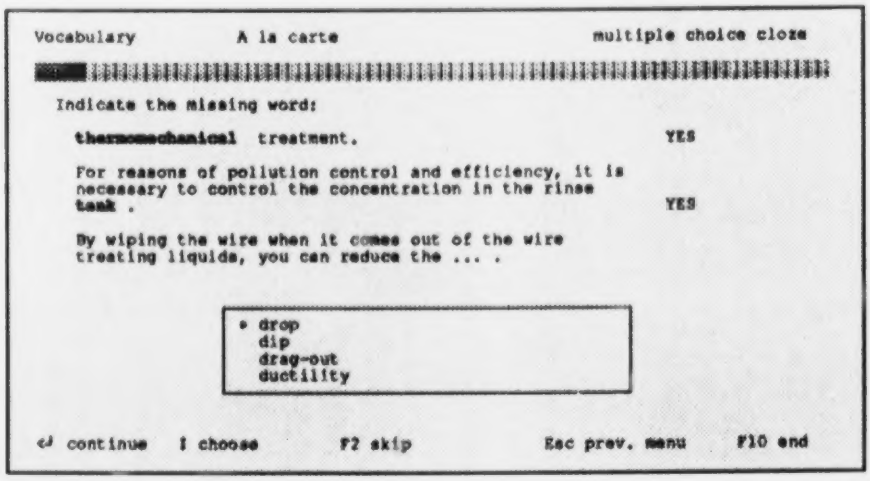


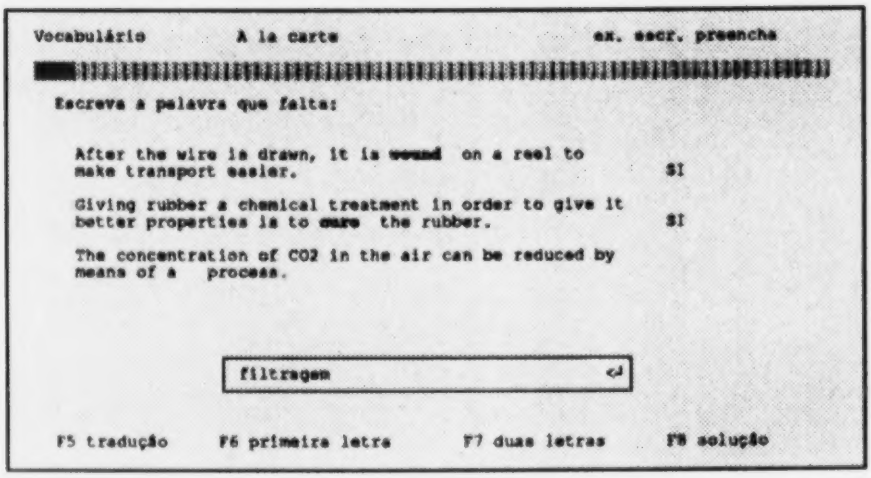
We can moreover add to this selection the proficiency level parameter by adjusting the level in the 'which lesson?' menu. If we concentrate on the less frequent or more difficult adjectives (levels 5 and 6), we immediately get a more restricted set of 14 adjectives. The possibility of combining selection criteria on an extensive database creates of course an endless range of possible selection combinations: verbs expressing treatments of wire, nouns designating names of substances, etc.

Every single selection made by the user can then be presented to him or her in range of possible *learning strategies*.



The 'which strategy?' entry lists the possibilities: *see selections* gives a quick overview of the items selected, *presentation* gives you the item with its translation and an STS context. Under *exercises* the user will find four possible varieties of exercises: gap-filling exercises with or without hint, translation exercises from source to target language and vice versa. Each of these exercises can then be made in a written, oral or multiple choice mode.


Vocabulary A la carte multiple choice cloze
Indicate the missing word:
thermomechanical treatment.
YES
For reasons of pollution control and efficiency, it is necessary to control the concentration in the rinse tank .
YES
By wiping the wire when it comes out of the wire treating liquids, you can reduce the ... .
• drop
dip
drag-out
ductility
↵ continue ↕ choose F2 skip Esc prev. menu F10 end


Vocabulário
A la carte
ex. secr. preencha
Escreva a palavra que falta:
After the wire is drawn, it is wound on a reel to make transport easier.
SI
Giving rubber a chemical treatment in order to give it better properties is to cure the rubber.
SI
The concentration of CO2 in the air can be reduced by means of a process.
filtragem
F5 tradução
F6 primeira letra
F7 duas letras
F8 solução

While making exercises, the user will get immediate feedback, automatic correction for minor mistakes (wrong articles, typing errors, etc.) and there is always help available (translation, article, first letters, etc.). This feedback and the help routines are no longer available when exercises are made in *test mode*. After the test one will get one's final score and the possibility to have an overview of one's mistakes.

4.2.3. Digestif

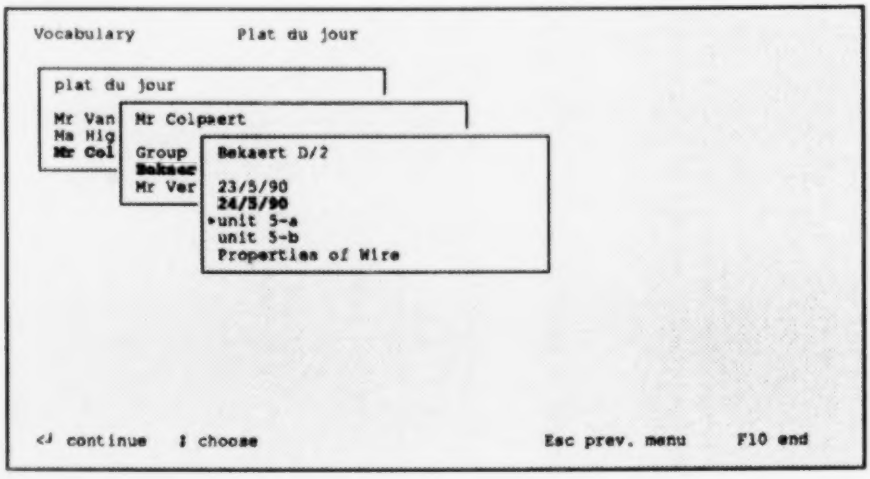
A second environment is called 'digestif'. In 'digestif' a student will, after a working session in 'à la carte' or 'plat du jour', find remedial material. That is to say, while exercises and tests are being made, the program keeps track of all occurring mistakes and enters them in the 'digestif' environment. The submenu of 'digestif' looks exactly as the strategies menu in 'à la carte': students can first have a look at their mistakes ('see selections'), study the items again

('presentation') and start a new exercise or test on the restricted subject matter.

While doing the 'digestif' the program will automatically generate a new digestif, based on mistakes made in the current digestif. In this way, the acquisition and error correction process is organized in a cyclical way: each time a new digestif is created, there will be less items to be redone by the students.

4.2.4. Plat du jour and trainer's entry

'Plat du jour' is meant for trainees who are monitored by language trainers. The environment offers the user a precise content which was defined in advance by the language trainer.



'Plat du jour' avoids that the trainee must make the rather complex 'à la carte' content and strategy selections by himself. Instead, the trainee will simply select 'plat du jour' and s/he will receive *an overview of the lessons available*, organized according to the language trainers using the program, according to the various

language groups a given trainer is working with, and according to the specific lessons planned for each group. A simple three times 'enter' will do for the selection of a training unit in 'plat du jour'. As soon as the trainee has selected his learning unit, it is composed from the database in a fraction of a second, exactly as the trainer has predefined it through using the separate teacher's entry.

The *teacher's entry* is meant to quickly define and store on disk the exact contents of a lesson for a particular class or student. It is in other words the way for a teacher to organize and prepare lessons in the 'plat du jour'. The teacher will first identify the unit by giving it a threefold identification, for example his own name, the identification of his group and the title of a given unit. This identification is the only thing the student will have to know in order to enter the predefined unit. The same menus as in 'à la carte' allow a teacher to define the unit. It takes an average of 20 seconds to do this, since the teacher does not have to type anything, s/he only has to indicate the precise limits within the potential and to determine the strategies s/he thinks appropriate for the contents and the target public selected. The units thus defined and stored on disk can be changed, expanded or erased, whenever the teacher wants to.

4.2.5. Buffet

Finally it might well be that a user would like to apply the program as a database for consultation rather than for learning motives. This can be done in the exploratory environment 'buffet'. Users can quickly consult the program's

database for checking the translation of a given word, for searching meaningful contexts for a particular word, etc.

5. Conclusions

Thanks to an efficiently organized collaboration between university and industrial partners, Calliope has been able to develop relevant and made-to-measure courseware materials for business and industrial purposes. Crucial aspects of such a development planning were a careful investigation of existing language needs in the company, the collection of company authentic data materials, an intelligent analysis and encoding of these materials for database use, carefully developed strategies and learning contexts and an externally versatile user interface which increases the multifunctionality of the developed program. As mentioned at the beginning of this article, this is not the complete Calliope story. A 'last but not least' stage is the implementation of the courseware in the existing training infrastructure of the industrial partner. Such an implementation, though facilitated by a careful project development in continuous and mutual consultancy, is a 'conditio sine qua non' for the efficient and effective use of the program in daily company training situations.

---

[1]Research Assistant of the National Foundation for Scientific Research (Belgium - NFWO)

[2]Computer Assisted Language Learning for Information, Organization and Production in Europe.

[3]Language for Specific (or Specialized) Purposes

[4]For more details on the database management system used in the Didascalia programs, see N. Palm and E. Van Elsen, *'The Formatting of Linguistic Content Bases: Meeting the Criteria of External Versatility.'* (in prep.).

[5]see also: J. Colpaert, W. Decoo and E. Van Elsen; *'Courseware for LSP (Language for Special Purposes): Creating Units at Will in a Selection System'*. In: Computers in Education; F. Lovis and E.D. Taggs [eds.], Elsevier Science Publishers B.V. (North-Holland) 1988, p. 553-558.

END

U.S. Dept. of Education

Office of Education
Research and
Improvement (OERI)

ERIC

Date Filmed

March 29, 1991